

# NOVEL STARCH SYNTHASE POLYNUCLEOTIDES AND THEIR USE IN THE PRODUCTION OF NEW STARCHES

George W. Singletary

Lan Zhou

## TECHNICAL FIELD

The present invention relates generally to plant molecular biology. More specifically, it relates to nucleic acids and methods for modulating their expression in plants.

## CROSS-REFERENCE TO RELATED APPLICATIONS

*divisional*  
This application claims priority under 35 USC 120 to U.S. Patent Application Serial No. 09/388,743 filed September 2, 1999, *now U.S. Patent 6,423,886*, the disclosure of which is incorporated herein by reference.

## BACKGROUND OF THE INVENTION

Starches are polymers of glucose molecules produced and stored only in the chloroplasts and amyloplasts of plants. Most of the starch produced in the world is used as food, but about one-third of the total production is employed for a variety of industrial purposes that take advantage of starch's unique properties. These properties (e.g. viscosity, gelatinization temperature) vary greatly with the plant source and affect the usefulness of the starch for food and nonfood products. (Sivak and Preiss, *Advances in Food and Nutrition Research*, Vol. 41. Academic Press, 1998, p. 163.) The properties of a starch are determined by the structure of its constituent granules and the polymers from which the granules are formed.

Molecular structure is, in turn, dependent largely upon the presence and activity of specific isozymes of starch synthase. Genes encoding the starch synthase enzymes are therefore of interest.

The presence or absence of specific starch synthase isozymes in tissues can have profound effects upon the nature of starch polymers and granules produced in plants, although the exact role of different synthase isozymes in defining the specificity of amylose, amylopectin and granule properties is still unclear. Recent  
5 research suggests that the contribution of a starch synthase isozyme is determined not only by intrinsic properties of the enzyme but also by interactions among various synthases, branching enzymes, and/or debranching enzymes. Thus, the genetic, environmental, and developmental backgrounds in which a particular starch synthase isozyme is expressed affect the role of that isozyme in dictating distinct features of  
10 starch. This implies that almost any manipulation of combinations of starch synthesizing enzymes may have surprising and potentially useful results. (Smith, A.M. *Current Opinion in Plant Biology* 2:223-229 (1999)).

When a mutation affecting a particular starch biosynthetic enzyme results in changes in the appearance of a seed, the resulting changes in starch structure may  
15 be subtle: for example, a slight decrease in the average chain length of amylopectin or a small increase in the proportion of amylose to amylopectin. (Sivak and Preiss, supra, p. 29.) However, even small changes in the molecular structure of a starch may have significant effects on its industrial utility. Identification of enzymatic changes and of the consequent modification of starch will result in enhanced diversity of starch  
20 functionalities for industrial purposes.

Certain genes encoding starch synthases have been identified and cloned, and modifications of starch content using these genes have been attempted or accomplished. See, for example, US Patent No. 5,824,798; WO Publication No. 96/15248; WO Publication No. 97/45545; WO Publication No. 97/26362; and WO  
25 Publication No. 97/20936.

Starch synthase enzymes utilize ADPglucose and/or UDPglucose in a polymerization reaction. The glycosyl portion of the substrate is transferred to, in most cases, preexisting maltooligosaccharides or polymers of  $\alpha$ -(1,4) or mixed  $\alpha$ -(1,4) and  $\alpha$ -(1,6) linkages. In other words, these enzymes are involved in the  
30 biosynthesis of amylose [ $\alpha$ -(1,4) polymer] and amylopectin [polymer of  $\alpha$ -(1,4)

branched with  $\alpha$ -(1-4,6) linkages], the primary types of molecules occurring in starch granules.

Starch composition varies with species and tissue. Maize starch stored in the endosperm is typically composed of 28% amylose and 72% amylopectin. Amylose of  
5 maize starch has an average degree of polymerization (number-average;  $DP_n$ ) of 960 and an average chain length (CL) of 305. Maize amylopectin, on the other hand, has a  $DP_n$  of  $10.2 \times 10^3$  and a CL of 22. (Morrison et al., "Starch" in Dey & Harborne, *Methods in Plant Biochemistry* (London, Academic Press, 1990), pp. 323-352).

The different structures of amylose and amylopectin confer distinctive  
10 properties to these polysaccharides. For example, the amylose fraction of starch will precipitate from an aqueous solution over time as the linear amylose molecules align themselves parallel to each other and become held together by hydrogen bonds. This precipitation is known as retrogradation. Amylopectin, on the other hand, does not readily form intermolecular complexes and is more stable in aqueous solutions.  
15 However, the crystalline nature of starch is due to the presence of amylopectin and not to amylose. This is contrary to the general principle that branching in a molecule is detrimental to crystallization. (Sivak and Preiss, supra, pp. 20-22) Indeed, the chain length of amylopectin is a basic factor in the determination of the crystalline type of the starch. (Hizukuri, S., *Carbohydrate Research* 141, pp. 295-306 (1985)).

20 The ease of isolation of starch granules for food and industrial purposes is affected by starch composition. In each bushel of maize processed, roughly 2.2 to 3.4 pounds of starch are unrecoverable. This represents a 6% to 8% loss of potential starch yield. Thus, a need exists for improved yield of starch from maize grain; alteration of granule size and/or density could improve this processing yield.

25 Starches with unusual, desirable functional properties may be currently available only in small quantities, or in plants or plant parts not commonly processed. Therefore, a need exists to develop plants, especially cereals, potato, or cassava, capable of synthesizing unusual starch for use in specific food and non-food industrial applications.

30

## SUMMARY OF THE INVENTION

It is an object of the present invention to provide nucleic acids and polypeptides relating to the biosynthesis of starch.

It is another object of the present invention to provide nucleic acids and polypeptides that can be used to identify proteins involved in starch biosynthesis.

It is another object of the present invention to provide antigenic fragments of the polypeptides of the present invention.

It is another object of the present invention to provide transgenic plants comprising the nucleic acids of the present invention.

It is another object of the present invention to provide methods for modulating, in a transgenic plant, the expression of the nucleic acids of the present invention.

It is another object of the present invention to provide a method for modulating the type and level of starch synthase in a plant.

It is another object of the present invention to provide a method for modulating the type and level of starch in a plant.

It is another object of the present invention to improve the industrial processing yield of starch from maize grain by altering starch granule size and/or density.

It is another object of the present invention to provide plants, especially maize, wheat, sorghum, barley, millet, rice, potato, or cassava plants, capable of synthesizing unusual starch with desirable functional properties for use in specific food and non-food industrial applications.

It is another object of the present invention to provide plants capable of synthesizing starch with altered granule sizes.

It is another object of the present invention to provide plants capable of synthesizing starch with altered degrees of crystallinity.

It is another object of the present invention to provide plants capable of synthesizing starch with altered densities.

It is another object of the present invention to provide plants capable of synthesizing starch with altered digestibilities.

It is another object of the present invention to provide plants capable of synthesizing starch with altered levels of covalently bound phosphate.

It is another object of the present invention to provide plants capable of synthesizing starch with altered patterns of branching and/or average chain lengths.

It is another object of the present invention to provide plants capable of synthesizing starch with altered temperatures of gelatinization.

5 It is another object of the present invention to provide plants capable of synthesizing starch with altered degrees of polymerization and retrogradation.

Therefore, in one aspect, the present invention relates to an isolated nucleic acid comprising a member selected from the group consisting of:

- 10 (a) a polynucleotide that encodes a polypeptide of SEQ ID NO: 2, 6, 10, 14, 18, 22, or 26;
- (b) a polynucleotide amplified from a Zea mays nucleic acid library using one or more primers selected from SEQ ID NOS: 3, 4, 7, 8, 11, 12, 15, 16, 19, 20, 23, 24, 27, and 28;
- 15 (c) a polynucleotide comprising at least 20 contiguous bases of SEQ ID NO: 1, 5, 9, 13, 17, 21, or 25;
- (d) a polynucleotide encoding a starch synthase from *Curcuma zedoaria*, *Maranta arundinacea*, *Canna edulis*, *Typha latifolia*, or *Tulipa fosteriana*;
- 20 (e) a polynucleotide having at least 73% sequence identity to SEQ ID NO: 1, 5, 9, 13, 17, 21, or 25, wherein the % sequence identity is based on the entire coding sequence and is determined by BLAST 2.0 using default parameters;
- 25 (f) a polynucleotide comprising at least 25 nucleotides in length which hybridizes under low stringency conditions to a polynucleotide having the sequence set forth in SEQ ID NO: 1, 5, 9, 13, 17, 21, or 25, wherein the conditions include a wash in 0.1X SSC at 60°C;
- (g) a polynucleotide comprising the coding sequence set forth in SEQ ID NO: 1, 5, 9, 13, 17, 21, or 25; and
- (h) a polynucleotide complementary to a polynucleotide of (a) through (g).

Other aspects of the present invention include expression cassettes  
30 comprising the nucleic acid operably linked to a promoter, host cells transfected with

the expression cassette, and transgenic plants and seeds comprising the expression cassette.

Also provided is an isolated protein comprising a member selected from the group consisting of:

- 5 (a) a polypeptide comprising at least 25 contiguous amino acids of SEQ ID NO: 2, 6, 10, 14, 18, 22, or 26;
- (b) a polypeptide which is a starch synthase protein from *Curcuma zedoaria*, *Maranta arundinacea*, *Canna edulis*, *Typha latifolia*, or *Tulipa fosteriana*;
- 10 (c) a polypeptide comprising at least 73% sequence identity to SEQ ID NO: 2, 6, 10, 14, 18, 22, or 26, wherein the % sequence identity is based on the entire sequence and is determined by BLAST 2.0 using default parameters;
- (d) a polypeptide encoded by a nucleic acid of claim 1; and
- 15 (e) a polypeptide characterized by SEQ ID NO: 2, 6, 10, 14, 18, 22, or 26.

In a further aspect, the present invention relates to a method of modulating expression of the nucleic acids in a plant, comprising the steps of:

- 20 (a) transforming a plant cell with an expression cassette comprising a nucleic acid of the present invention operably linked to a promoter in sense or antisense orientation;
- (b) growing the plant cell under plant growing conditions to produce a regenerated plant capable of expressing the nucleic acid for a time sufficient to modulate expression of the nucleic acids in the plant compared to a corresponding non-transformed plant.

25 Expression of the nucleic acids encoding the proteins of the present invention can be increased or decreased relative to a non-transformed control plant.

### DETAILED DESCRIPTION OF THE INVENTION

30 The genes presently claimed encode starch synthases derived from plant species known to contain starch with unusual properties. Expression of these genes in other species, such as maize, wheat, sorghum, barley, millet, rice, potato, or

cassava, is expected to modify the structure of starch polymers and granules in those species. The result will be improved industrial isolation of starch and the production of starch with novel functional properties currently unobtainable in agronomic crops.

Starch granules of *Curcuma zedoaria* have an unusual elongated morphology, are about 70 $\mu$  in length and 14 $\mu$  in thickness, and have a temperature of gelatinization of about 83 °C. (Jane, J., et al., *Starch* 46:121-129 (1994); Reichert, E.T., *The Differentiation and Specificity of Starches in Relation to Genera, Species, Etc.* Carnegie Institution, Washington, D.C. (1913)). In contrast, maize granules have a round, polygonal form; range in size from about 2 $\mu$  to 30 $\mu$ , with an average diameter of 10 $\mu$  (Swinkels, J.J.M. *Starch* 37:1-5 (1985)); and gelatinize at about 65°C.

Tulip starch contains approximately 27% amylose (Deatherage, W.L., et al. *Trans. Am. Assoc. Cereal Chem.* 13:31-42 (1955)) and has a granule size of roughly 6 $\mu$  to 50 $\mu$ , with an average diameter of 35 $\mu$ . Most notably, tulip starch has a gelatinization temperature of about 50 °C, (Reichert, E.T., *supra*) and its amylopectin has a chain length longer than that of maize starch. Tulip starch also has a pattern of branching different from that of maize starch. (Hizukuri, S., *supra*).

Expression of the isolated starch synthase genes of the present application in plants, especially maize, wheat, sorghum, barley, millet, rice, potato, or cassava plants, could allow the production of starch with unique functional properties not currently available. For example, it is expected that expression of the present starch synthase genes in maize could:

- (1) Lead to the biosynthesis of a population of starch granules that are, on average, larger than the granules which occur in normal maize. For example, the starch granules could range in size from 35 $\mu$  to 100 $\mu$ , from 40 $\mu$  to 100 $\mu$ , or from 50 $\mu$  to 100 $\mu$ .
- (2) Lead to the biosynthesis of starch granules that have greater or lesser degrees of crystallinity and gelatinization temperatures and, in turn, elevated or reduced densities. Greater density could facilitate recovery of starch in wet milling, and increased gelatinization temperatures could

allow the use of higher temperatures (i.e. reaction rates) in chemical modification of starch. Starch of decreased density or gelatinization temperature could be more digestible as a food material or in feedstuffs and could offer process efficiency in the conversion of intact granules into gelatinized granule fragments.

- (3) Lead to the biosynthesis of starch granules that have a greater degree of crystallinity and display reduced digestibility. Such a starch could have application in low-calorie foods and/or sport drinks which help maintain lower, stable glycemic indices.
- (4) Result in higher levels of covalently bound phosphate by way of a synergistic interaction with the native mechanism which normally allows only low levels of phosphorylation.
- (5) Produce altered patterns of branching and modified average chain lengths in amylopectin which could produce starch that has a reduced temperature of gelatinization.
- (6) Produce amylose of lower or higher degree of polymerization than in normal maize starch and lead to altered rates of retrogradation.

## DEFINITIONS

The term "isolated" refers to material, such as a nucleic acid or a protein, which is: (1) substantially or essentially free from components which normally accompany or interact with it as found in its natural environment. The isolated material optionally comprises material not found with the material in its natural environment; or (2) if the material is in its natural environment, the material has been synthetically altered or synthetically produced by deliberate human intervention and/or placed at a different location within the cell. The synthetic alteration or creation of the material can be performed on the material within or apart from its natural state. For example, a naturally occurring nucleic acid becomes an isolated nucleic acid if it is altered or produced by non-natural, synthetic methods, or if it is transcribed from DNA which has been altered or produced by non-natural, synthetic methods. The isolated nucleic acid may also be produced by the synthetic re-



arrangement ("shuffling") of a part or parts of one or more allelic forms of the gene of interest. Likewise, a naturally-occurring nucleic acid (e.g., a promoter) becomes isolated if it is introduced to a different locus of the genome. Nucleic acids which are "isolated," as defined herein, are also referred to as "heterologous" nucleic acids.

5 As used herein, "polypeptide" means proteins, protein fragments, modified proteins, amino acid sequences and synthetic amino acid sequences. The polypeptide can be glycosylated or not.

As used herein, "plant" includes but is not limited to plant cells, plant tissue and plant seeds.

10 As used herein, "promoter" includes reference to a region of DNA upstream from the start of transcription and involved in recognition and binding of RNA polymerase and other proteins to initiate transcription.

By "fragment" is intended a portion of the nucleotide sequence or a portion of the amino acid sequence and hence protein encoded thereby. Preferably fragments of a nucleotide sequence may encode protein fragments that retain the biological activity of the native nucleic acid. However, fragments of a nucleotide sequence which are useful as hybridization probes generally do not encode fragment proteins retaining biological activity. Fragments of a nucleotide sequence are generally greater than 10 nucleotides, preferably at least 20 nucleotides and up to the entire nucleotide sequence encoding the proteins of the invention. Generally probes are less than 1000 nucleotides and preferably less than 500 nucleotides. Fragments of the invention include antisense sequences used to decrease expression of the inventive nucleic acids. Such antisense fragments may vary in length ranging from at least about 20 nucleotides, about 50 nucleotides, about 100 nucleotides, up to and including the entire coding sequence.

25 By "variants" is intended substantially similar sequences. Generally, nucleic acid sequence variants of the invention will have at least 50%, 60%, 70%, or preferably 80%, more preferably at least 90% and most preferably at least 95% sequence identity to the native nucleotide sequence.

Generally, polypeptide sequence variants of the invention will have at least about 55%, 60%, 70%, 80%, or preferably at least about 90% and more preferably at least about 95% sequence identity to the native protein.

As used herein, "sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences includes reference to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window. An indication that two peptide sequences are substantially identical is that one peptide is immunologically reactive with antibodies raised against the second peptide. A polypeptide is substantially identical to a second polypeptide, for example, where the two polypeptides differ only by a conservative substitution.

Methods of alignment of sequences for comparison are well-known in the art. For purposes of defining the present invention, the BLAST 2.0 suite of programs using default parameters is used. Altschul *et al.*, *Nucleic Acids Res.* 25:3389-3402 (1997). Software for performing BLAST analyses is publicly available, e.g., through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>).

By "functional equivalent, variant, or derivative" is intended a sequence that produces a protein having substantially the same biological effect as the native protein of interest.

## NUCLEIC ACIDS

The isolated nucleic acids of the present invention can be made using (a) standard recombinant methods, (b) synthetic techniques, or (c) combinations thereof. The isolated nucleic acid may also be produced by the synthetic re-arrangement ("shuffling") of a part or parts of one or more allelic forms of the gene of interest. Sequence shuffling is described in PCT publication No. WO97/20078. See also, Zhang, J. H., et al., *Proc. Natl. Acad. Sci. USA* 94:4504-4509 (1997).

In some embodiments, the polynucleotides of the present invention will be cloned, amplified, or otherwise constructed from a plant, including potato, cassava, maize, soybean, sunflower, canola, alfalfa, cotton, *Arabidopsis*, tomato, pepper, apple, spinach, and lettuce. In preferred embodiments the plant is *Curcuma zedoaria*

or other species of the genus *Curcuma*; *Tulipa fosteriana* or other species of the genus *Tulipa*; *Maranta arundinacea* or other species of the genus *Maranta*; *Canna edulis* or other species of the genus *Canna*; *Typha latifolia* or other species of the genus *Typha*. Source tissues for isolated polynucleotides are preferably starchy plant parts, including tuber, root, and rhizome.

Functional fragments included in the invention can be obtained using primers that selectively hybridize under stringent conditions. Primers are generally at least 12 bases in length and can be as high as 200 bases, but will generally be from 15 to 75, preferably from 15 to 50. Functional fragments can be identified using a variety of techniques such as restriction analysis, Southern analysis, primer extension analysis, and DNA sequence analysis.

The present invention includes a plurality of polynucleotides that encode for the identical amino acid sequence. The degeneracy of the genetic code allows for such "silent variations" which can be used, for example, to selectively hybridize and detect allelic variants of polynucleotides of the present invention. Additionally, the present invention includes isolated nucleic acids comprising allelic variants. The term "allele" as used herein refers to a related nucleic acid of the same gene.

Variants of nucleic acids included in the invention can be obtained, for example, by oligonucleotide-directed mutagenesis, linker-scanning mutagenesis, mutagenesis using the polymerase chain reaction, and the like. See, for example, Ausubel, pages 8.0.3 - 8.5.9. Also, see generally, McPherson (ed.), *DIRECTED MUTAGENESIS: A Practical approach*, (IRL Press, 1991). Thus, the present invention also encompasses DNA molecules comprising nucleotide sequences that have substantial sequence similarity with the inventive sequences.

Variants included in the invention may contain individual substitutions, deletions or additions to the nucleic acid or polypeptide sequences. Such changes will alter, add or delete a single amino acid or a small percentage of amino acids in the encoded sequence. Variants are referred to as "conservatively modified variants" where the alteration results in the substitution of an amino acid with a chemically similar amino acid. When the nucleic acid is prepared or altered synthetically, advantage can be taken of known codon preferences of the intended host.

The present invention also includes the use of 5' and/or 3' UTR regions for modulation of translation of heterologous coding sequences. Positive sequence motifs include translational initiation consensus sequences (Kozak, *Nucleic Acids Res.* 15:8125 (1987)) and the 7-methylguanosine cap structure (Drummond et al.,  
5 *Nucleic Acids Res.* 13:7375 (1985)). Negative elements include stable intramolecular 5' UTR stem-loop structures (Muesing et al., *Cell* 48:691 (1987)) and AUG sequences or short open reading frames preceded by an appropriate AUG in the 5' UTR (Kozak, *supra*, Rao et al., *Mol. and Cell. Biol.* 8:284 (1988)).

Further, the polypeptide-encoding segments of the polynucleotides of the  
10 present invention can be modified to alter codon usage. Altered codon usage can be employed to alter translational efficiency and/or to optimize the coding sequence for expression in a desired host or to optimize the codon usage in a heterologous sequence for expression in maize. Codon usage in the coding regions of the polynucleotides of the present invention can be analyzed statistically using  
15 commercially available software packages such as "Codon Preference" available from the University of Wisconsin Genetics Computer Group (see Devereaux et al., *Nucleic Acids Res.* 12:387-395 (1984)) or MacVector 4.1 (Eastman Kodak Co., New Haven, Conn.).

For example, the inventive nucleic acids can be optimized for enhanced  
20 expression in organisms of interest. See, for example, EPA0359472; WO91/16432; Perlak et al., *Proc. Natl. Acad. Sci. USA* 88:3324-3328 (1991); and Murray et al., *Nucleic Acids Res.* 17:477-498 (1989). In this manner, the genes can be synthesized utilizing species-preferred codons. See, for example, Murray et al., *Nucleic Acids Res.* 17:477-498 (1989), the disclosure of which is incorporated herein by reference.

25 The present invention provides subsequences comprising isolated nucleic acids containing at least 16 contiguous bases of the inventive sequences. For example the isolated nucleic acid includes those comprising at least 20, 25, 30, 40, 50, 60, 75 or 100 contiguous nucleotides of the inventive sequences. Subsequences of the isolated nucleic acid can be used to modulate or detect gene expression by  
30 introducing into the subsequences compounds which bind, intercalate, cleave and/or crosslink to nucleic acids.

The nucleic acids of the invention may conveniently comprise a multi-cloning site comprising one or more endonuclease restriction sites inserted into the nucleic acid to aid in isolation of the polynucleotide. Also, translatable sequences may be inserted to aid in the isolation of the translated polynucleotide of the present invention. For example, a hexa-histidine marker sequence provides a convenient means to purify the proteins of the present invention.

A polynucleotide of the present invention can be attached to a vector, adapter, promoter, transit peptide or linker for cloning and/or expression of a polynucleotide of the present invention. Additional sequences may be added to such cloning and/or expression sequences to optimize their function in cloning and/or expression, to aid in isolation of the polynucleotide, or to improve the introduction of the polynucleotide into a cell. Use of cloning vectors, expression vectors, adapters, and linkers is well known and extensively described in the art. For a description of such nucleic acids see, for example, Stratagene Cloning Systems, Catalogs 1995, 1996, 1997 (La Jolla, CA); and, Amersham Life Sciences, Inc, Catalog '97 (Arlington Heights, IL).

The isolated nucleic acid compositions of this invention, such as RNA, cDNA, genomic DNA, or a hybrid thereof, can be obtained from plant biological sources using any number of cloning methodologies known to those of skill in the art. In some embodiments, oligonucleotide probes that selectively hybridize, under stringent conditions, to the polynucleotides of the present invention are used to identify the desired sequence in a cDNA or genomic DNA library.

Exemplary total RNA and mRNA isolation protocols are described in *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997); and, *Current Protocols in Molecular Biology*, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). Total RNA and mRNA isolation kits are commercially available from vendors such as Stratagene (La Jolla, CA), Clontech (Palo Alto, CA), Pharmacia (Piscataway, NJ), and 5'-3' (Paoli, PA). See also, U.S. Patent Nos. 5,614,391; and, 5,459,253.

Typical cDNA synthesis protocols are well known to the skilled artisan and are described in such standard references as: *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997); and, *Current Protocols in*

*Molecular Biology*, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). cDNA synthesis kits are available from a variety of commercial vendors such as Stratagene or Pharmacia.

An exemplary method of constructing a greater than 95% pure full-length  
5 cDNA library is described by Carninci et al., *Genomics* 37:327-336 (1996). Other methods for producing full-length libraries are known in the art. See, e.g., Edery et al., *Mol. Cell Biol.* 15(6):3363-3371 (1995); and, PCT Application WO 96/34981.

It is often convenient to normalize a cDNA library to create a library in which  
each clone is more equally represented. A number of approaches to normalize cDNA  
10 libraries are known in the art. Construction of normalized libraries is described in Ko, *Nucl. Acids. Res.* 18(19):5705-5711 (1990); Patanjali et al., *Proc. Natl. Acad. USA* 88:1943-1947 (1991); U.S. Patents 5,482,685 and 5,637,685; and Soares et al., *Proc. Natl. Acad. Sci. USA* 91:9228-9232 (1994).

Subtracted cDNA libraries are another means to increase the proportion of  
15 less abundant cDNA species. See Foote et al. in, *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag, Berlin (1997); Kho and Zarbl, *Technique*, 3(2):58-63 (1991); Sive and St. John, *Nucl. Acids Res.* 16(22):10937 (1988); *Current Protocols in Molecular Biology*, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995); and, Swaroop et al., *Nucl. Acids*  
20 *Res.* 19(8):1954 (1991). cDNA subtraction kits are commercially available. See, e.g., PCR-Select (Clontech).

To construct genomic libraries, large segments of genomic DNA are generated  
by random fragmentation. Examples of appropriate molecular biological techniques  
and instructions are found in Sambrook et al., *Molecular Cloning: A Laboratory*  
25 *Manual*, 2nd Ed., Cold Spring Harbor Laboratory Vols. 1-3 (1989), *Methods in Enzymology*, Vol. 152: *Guide to Molecular Cloning Techniques*, Berger and Kimmel, Eds., San Diego: Academic Press, Inc. (1987), *Current Protocols in Molecular Biology*, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995); *Plant Molecular Biology: A Laboratory Manual*, Clark, Ed., Springer-Verlag,  
30 Berlin (1997). Kits for construction of genomic libraries are also commercially available.

The cDNA or genomic library can be screened using a probe based upon the sequence of a nucleic acid of the present invention such as those disclosed herein. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different plant species. Those of skill in the art will appreciate that various degrees of stringency of hybridization can be employed in the assay; and either the hybridization or the wash medium can be stringent. The degree of stringency can be controlled by temperature, ionic strength, pH and the presence of a partially denaturing solvent such as formamide.

Typically, stringent hybridization conditions will be those in which the salt concentration is less than about 1.5 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Typically the hybridization will be conducted for about 4 to about 12 hours.

Preferably the hybridization is conducted under low stringency conditions which include hybridization with a buffer solution of 30 % formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulfate) at 37°C, and a wash in 1X to 2X SSC (20X SSC = 3.0 M NaCl/0.3 M trisodium citrate) at 50°C. More preferably the hybridization is conducted under moderate stringency conditions which include hybridization in 40 % formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55°C. Most preferably the hybridization is conducted under high stringency conditions which include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1X SSC at 60°C. Typically the hybridization will be conducted for about 4 to about 12 hours.

An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); and *Current Protocols in Molecular Biology*, Chapter 2, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). Often, cDNA libraries will be normalized to increase the representation of relatively rare cDNAs.

The nucleic acids of the invention can be amplified from nucleic acid samples using amplification techniques. For instance, polymerase chain reaction (PCR) technology can be used to amplify the sequences of polynucleotides of the present invention and related genes directly from genomic DNA or cDNA libraries. PCR and other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes.

Examples of techniques useful for *in vitro* amplification methods are found in Berger, Sambrook, and Ausubel, as well as Mullis *et al.*, U.S. Patent No. 4,683,202 (1987); and, *PCR Protocols A Guide to Methods and Applications*, Innis *et al.*, Eds., Academic Press Inc., San Diego, CA (1990). Commercially available kits for genomic PCR amplification are known in the art. See, e.g., Advantage-GC Genomic PCR Kit (Clontech). The T4 gene 32 protein (Boehringer Mannheim) can be used to improve yield of long PCR products.

PCR-based screening methods have also been described. Wilfinger *et al.* describe a PCR-based method in which the longest cDNA is identified in the first step so that incomplete clones can be eliminated from study. *BioTechniques*, 22(3):481-486 (1997).

In one aspect of the invention, nucleic acids can be amplified from a plant nucleic acid library. The nucleic acid library may be a cDNA library, a genomic library, or a library generally constructed from nuclear transcripts at any stage of intron processing. Libraries can be made from a variety of plant tissues.

Alternatively, the sequences of the invention can be used to isolate corresponding sequences in other organisms, particularly other plants, more particularly, other monocots. In this manner, methods such as PCR, hybridization, and the like can be used to identify such sequences having substantial sequence similarity to the sequences of the invention. See, for example, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York). and Innis *et al.* (1990), *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York). Coding sequences



isolated based on their sequence identity to the entire inventive coding sequences set forth herein or to fragments thereof are encompassed by the present invention.

The isolated nucleic acids of the present invention can also be prepared by direct chemical synthesis by methods such as the phosphotriester method of Narang et al., *Meth. Enzymol.* 68:90-99 (1979); the phosphodiester method of Brown et al., *Meth. Enzymol.* 68:109-151 (1979); the diethylphosphoramidite method of Beaucage et al., *Tetra. Lett.* 22:1859-1862 (1981); the solid phase phosphoramidite triester method described by Beaucage and Caruthers, *Tetra. Letts.* 22(20):1859-1862 (1981), e.g., using an automated synthesizer, e.g., as described in Needham-VanDevanter et al., *Nucleic Acids Res.*, 12:6159-6168 (1984); and, the solid support method of U.S. Patent No. 4,458,066. Chemical synthesis generally produces a single stranded oligonucleotide. This may be converted into double stranded DNA by hybridization with a complementary sequence, or by polymerization with a DNA polymerase using the single strand as a template. One of skill will recognize that while chemical synthesis of DNA is limited to sequences of about 100 bases, longer sequences may be obtained by the ligation of shorter sequences.

## EXPRESSION CASSETTES

In another embodiment expression cassettes comprising isolated nucleic acids of the present invention are provided. An expression cassette will typically comprise a polynucleotide of the present invention operably linked to transcriptional initiation regulatory sequences which will direct the transcription of the polynucleotide in the intended host cell, such as tissues of a transformed plant.

The construction of expression cassettes that can be employed in conjunction with the present invention is well known to those of skill in the art in light of the present disclosure. See, e.g., Sambrook et al.; *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor, New York; (1989); Gelvin et al.; *Plant Molecular Biology Manual*; (1990); *Plant Biotechnology: Commercial Prospects and Problems*, eds. Prakash et al.; Oxford & IBH Publishing Co.; New Delhi, India; (1993); and Heslot et al.; *Molecular Biology and Genetic Engineering of Yeasts*; CRC Press, Inc., USA; (1992); each incorporated herein in its entirety by reference.

For example, plant expression vectors may include (1) a cloned plant nucleic acid under the transcriptional control of 5' and 3' regulatory sequences and (2) a dominant selectable marker. Such plant expression vectors may also contain, if desired, a promoter regulatory region (e.g., one conferring inducible, constitutive, environmentally- or developmentally-regulated, or cell- or tissue-specific/selective expression), a transcription initiation start site, a ribosome binding site, an RNA processing signal, a transcription termination site, and/or a polyadenylation signal.

Constitutive, tissue-preferred or inducible promoters can be employed. Examples of constitutive promoters include the cauliflower mosaic virus (CaMV) 35S transcription initiation region, the 1'- or 2'- promoter derived from T-DNA of *Agrobacterium tumefaciens*, the ubiquitin 1 promoter, the Smas promoter, the cinnamyl alcohol dehydrogenase promoter (U.S. Patent No. 5,683,439), the *Nos* promoter, the pEmu promoter, the rubisco promoter, the GRP1-8 promoter and other transcription initiation regions from various plant genes known to those of skill.

Examples of inducible promoters are the *Adh1* promoter which is inducible by hypoxia or cold stress, the *Hsp70* promoter which is inducible by heat stress, and the *PPDK* promoter which is inducible by light. Also useful are promoters which are chemically inducible.

Examples of promoters under developmental control include promoters that initiate transcription preferentially in certain tissues, such as leaves, roots, fruit, seeds, or flowers. An exemplary promoter is the anther specific promoter 5126 (U.S. Patent Nos. 5,689,049 and 5,689,051). Examples of seed-preferred promoters include, but are not limited to, 27 kD gamma zein promoter and waxy promote, Boronat,A., Martinez,M.C., Reina,M., Puigdomenech, P. and Palau,J.; Isolation and sequencing of a 28 kD glutelin-2 gene from maize: Common elements in the 5' flanking regions among zein and glutelin genes; *Plant Sci.* 47, 95-102 (1986) and Reina,M., Ponte,I., Guillen,P., Boronat,A. and Palau,J., Sequence analysis of a genomic clone encoding a Zc2 protein from *Zea mays* W64 A, *Nucleic Acids Res.* 18 (21), 6426 (1990). See the following site relating to the waxy promoter: Kloesgen, R.B., Gierl, A., Schwarz-Sommer, ZS. and Saedler,H., Molecular analysis of the waxy locus of *Zea mays*, *Mol. Gen. Genet.* 203, 237-244 (1986). Promoters that express in

the embryo, pericarp, and endosperm are disclosed in US applications Ser. Nos. 60/097,233 filed August 20, 1998 and 60/098,230 filed August 28, 1998. The disclosures of each of these are incorporated herein by reference in their entirety.

5 Either heterologous or non-heterologous (i.e., endogenous) promoters can be employed to direct expression of the nucleic acids of the present invention. These promoters can also be used, for example, in expression cassettes to drive expression of antisense nucleic acids to reduce, increase, or alter concentration and/or composition of the proteins of the present invention in a desired tissue.

10 If polypeptide expression is desired, it is generally desirable to include a polyadenylation region at the 3'-end of a polynucleotide coding region. The polyadenylation region can be derived from the natural gene, from a variety of other plant genes, or from T-DNA. The 3' end sequence to be added can be derived from, for example, the nopaline synthase or octopine synthase genes, or alternatively from another plant gene, or less preferably from any other eukaryotic gene.

15 An intron sequence can be added to the 5' untranslated region or the coding sequence of the partial coding sequence to increase the amount of the mature message that accumulates. See for example Buchman and Berg, *Mol. Cell Biol.* 8:4395-4405 (1988); Callis et al., *Genes Dev.* 1:1183-1200 (1987). Use of maize introns Adh1-S intron 1, 2, and 6, the Bronze-1 intron are known in the art. See  
20 generally, *The Maize Handbook*, Chapter 116, Freeling and Walbot, Eds., Springer, New York (1994).

The vector comprising the sequences from a polynucleotide of the present invention will typically comprise a marker gene which confers a selectable phenotype on plant cells. Usually, the selectable marker gene will encode antibiotic or herbicide  
25 resistance. Suitable genes include those coding for resistance to the antibiotic spectinomycin or streptomycin (e.g., the aada gene), the streptomycin phosphotransferase (SPT) gene coding for streptomycin resistance, the neomycin phosphotransferase (NPTII) gene encoding kanamycin or geneticin resistance, the hygromycin phosphotransferase (HPT) gene coding for hygromycin resistance.

30 Suitable genes coding for resistance to herbicides include those which act to inhibit the action of acetolactate synthase (ALS), in particular the sulfonylurea-type

herbicides (e.g., the acetolactate synthase (ALS) gene containing mutations leading to such resistance in particular the S4 and/or Hra mutations), those which act to inhibit action of glutamine synthase, such as phosphinothricin or basta (e.g., the *bar* gene), or other such genes known in the art. The *bar* gene encodes resistance to the herbicide basta and the ALS gene encodes resistance to the herbicide chlorsulfuron.

Typical vectors useful for expression of nucleic acids in higher plants are well known in the art and include vectors derived from the tumor-inducing (Ti) plasmid of *Agrobacterium tumefaciens* described by Rogers *et al.*, Meth. In Enzymol. 153:253-277 (1987). Exemplary *A. tumefaciens* vectors useful herein are plasmids pKYLX6 and pKYLX7 of Schardl *et al.*, Gene 61:1-11 (1987) and Berger *et al.*, Proc. Natl. Acad. Sci. USA 86:8402-8406 (1989). Another useful vector herein is plasmid pBI101.2 that is available from Clontech Laboratories, Inc. (Palo Alto, CA).

A variety of plant viruses that can be employed as vectors are known in the art and include cauliflower mosaic virus (CaMV), geminivirus, brome mosaic virus, and tobacco mosaic virus.

A polynucleotide of the present invention can be expressed in either sense or anti-sense orientation as desired. In plant cells, it has been shown that antisense RNA inhibits gene expression by preventing the accumulation of mRNA which encodes the enzyme of interest, see, e.g., Sheehy *et al.*, Proc. Nat'l. Acad. Sci. USA 85:8805-8809 (1988); and Hiatt *et al.*, U.S. Patent No. 4,801,340.

Another method of suppression is sense suppression. Introduction of nucleic acid configured in the sense orientation has been shown to be an effective means by which to block the transcription of target genes. For an example of the use of this method to modulate expression of endogenous genes see, Napoli *et al.*, The Plant Cell 2:279-289 (1990) and U.S. Patent No. 5,034,323.

A method of down-regulation of the protein involves using PEST sequences that provide a target for degradation of the protein.

Catalytic RNA molecules or ribozymes can also be used to inhibit expression of plant genes. The inclusion of ribozyme sequences within antisense RNAs confers RNA-cleaving activity upon them, thereby increasing the activity of the constructs.

The design and use of target RNA-specific ribozymes is described in Haseloff et al., *Nature* 334:585-591 (1988).

A variety of cross-linking agents, alkylating agents and radical generating species as pendant groups on polynucleotides of the present invention can be used to bind, label, detect, and/or cleave nucleic acids. For example, Vlassov, V. V., et al., *Nucleic Acids Res* (1986) 14:4065-4076, describe covalent bonding of a single-stranded DNA fragment with alkylating derivatives of nucleotides complementary to target sequences. A report of similar work by the same group is that by Knorre, D. G., et al., *Biochimie* (1985) 67:785-789. Iverson and Dervan also showed sequence-specific cleavage of single-stranded DNA mediated by incorporation of a modified nucleotide which was capable of activating cleavage (*J. Am. Chem. Soc.* (1987) 109:1241-1243). Meyer, R. B., et al., *J. Am. Chem. Soc.* (1989) 111:8517-8519, effect covalent crosslinking to a target nucleotide using an alkylating agent complementary to the single-stranded target nucleotide sequence. A photoactivated crosslinking to single-stranded oligonucleotides mediated by psoralen was disclosed by Lee, B. L., et al., *Biochemistry* (1988) 27:3197-3203. Use of crosslinking in triple-helix forming probes was also disclosed by Home et al., *J. Am. Chem. Soc.* (1990) 112:2435-2437. Use of N<sup>4</sup>, N<sup>4</sup>-ethanocytosine as an alkylating agent to crosslink to single-stranded oligonucleotides has also been described by Webb and Matteucci, *J. Am. Chem. Soc.* (1986) 108:2764-2765; *Nucleic Acids Res.* (1986) 14:7661-7674; Feteritz et al., *J. Am. Chem. Soc.* 113:4000 (1991). Various compounds to bind, detect, label, and/or cleave nucleic acids are known in the art. See, for example, U.S. Patent Nos. 5,543,507; 5,672,593; 5,484,908; 5,256,648; and, 5,681,941.

## PROTEINS

Proteins of the present invention include proteins derived from the native protein by deletion (so-called truncation), addition or substitution of one or more amino acids at one or more sites in the native protein. Such variants may result from, for example, genetic polymorphism or from human manipulation. Methods for such manipulations are generally known in the art.

For example, amino acid sequence variants of the polypeptide can be prepared by mutations in the cloned DNA sequence encoding the native protein of interest. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (MacMillan Publishing Company, New York); Kunkel (1985) *Proc. Natl. Acad. Sci. USA* 82:488-492; Kunkel et al. (1987) *Methods Enzymol.* 154:367-382; Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, New York); U.S. Patent No. 4,873,192; and the references cited therein; herein incorporated by reference. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff et al. (1978) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, D.C.), herein incorporated by reference. Conservative substitutions, such as exchanging one amino acid with another having similar properties, may be preferred.

In constructing variants of the proteins of interest, modifications to the nucleotide sequences encoding the variants will be made such that variants continue to possess the desired activity. Obviously, any mutations made in the DNA encoding the variant protein must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure. See EP Patent Application Publication No. 0075444 B1.

The isolated proteins of the present invention include a polypeptide comprising at least 23 contiguous amino acids encoded by any one of the nucleic acids of the present invention, or polypeptides which are conservatively modified variants thereof. The proteins of the present invention or variants thereof can comprise any number of contiguous amino acid residues from a polypeptide of the present invention, wherein that number is selected from the group of integers consisting of from 23 to the number of residues in a full-length polypeptide of the present invention. Optionally, this subsequence of contiguous amino acids is at least 25, 30, 35, or 40 amino acids in length, often at least 50, 60, 70, 80, or 90 amino acids in length.

The present invention includes catalytically active polypeptides (i.e., enzymes). Catalytically active polypeptides will generally have a specific activity of at least 20%,

30%, or 40%, and preferably at least 50%, 60%, or 70%, and most preferably at least 80%, 90%, or 95% that of the native (non-synthetic), endogenous polypeptide.

Further, the substrate specificity ( $k_{cat}/K_m$ ) is optionally substantially similar to the native (non-synthetic), endogenous polypeptide. Typically,  $k_{cat}/K_m$  will be at least  
5 30%, 40%, or 50%, that of the native (non-synthetic), endogenous polypeptide; and more preferably at least 60%, 70%, 80%, or 90%. Methods of assaying and quantifying measures of enzymatic activity and substrate specificity ( $k_{cat}/K_m$ ), are well known to those of skill in the art.

The present invention includes modifications that can be made to an inventive  
10 protein without diminishing its biological activity. Some modifications may be made to facilitate the cloning, expression, or incorporation of the targeting molecule into a fusion protein. Such modifications are well known to those of skill in the art and include, for example, a methionine added at the amino terminus to provide an initiation site, or additional amino acids (e.g., poly His) placed on either terminus to  
15 create conveniently located restriction sites or termination codons or purification sequences.

A protein of the present invention can be expressed in a recombinantly engineered cell such as bacteria, yeast, insect, mammalian, or preferably plant cells. The cells produce the protein in a non-natural condition (e.g., in quantity,  
20 composition, location, and/or time), because they have been genetically altered through human intervention to do so.

Typically, an intermediate host cell will be used in the practice of this invention to increase the copy number of the cloning vector. With an increased copy number, the vector containing the nucleic acid of interest can be isolated in significant  
25 quantities for introduction into the desired plant cells.

Host cells that can be used in the practice of this invention include prokaryotes, including bacterial hosts such as *Eschericia coli*, *Salmonella typhimurium*, and *Serratia marcescens*. Eukaryotic hosts such as yeast or filamentous fungi may also be used in this invention. It is preferred to use plant  
30 promoters that do not cause expression of the polypeptide in bacteria.

Commonly used prokaryotic control sequences include promoters such as the beta lactamase (penicillinase) and lactose (lac) promoter systems (Chang et al., *Nature* 198:1056 (1977)), the tryptophan (trp) promoter system (Goeddel et al., *Nucleic Acids Res.* 8:4057 (1980)) and the lambda derived P L promoter and N-gene ribosome binding site (Shimatake et al., *Nature* 292:128 (1981)). The inclusion of selection markers in DNA vectors transfected in *E. coli* is also useful. Examples of such markers include genes specifying resistance to ampicillin, tetracycline, or chloramphenicol.

The vector is selected to allow introduction into the appropriate host cell. Bacterial vectors are typically of plasmid or phage origin. Expression systems for expressing a protein of the present invention are available using *Bacillus sp.* and *Salmonella* (Palva et al., *Gene* 22:229-235 (1983); Mosbach et al., *Nature* 302:543-545 (1983)).

Synthesis of heterologous proteins in yeast is well known. See Sherman, F., et al., *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory (1982). Two widely utilized yeast for production of eukaryotic proteins are *Saccharomyces cerevisiae* and *Pichia pastoris*. Vectors, strains, and protocols for expression in *Saccharomyces* and *Pichia* are known in the art and available from commercial suppliers (e.g., Invitrogen). Suitable vectors usually have expression control sequences, such as promoters, including 3-phosphoglycerate kinase or alcohol oxidase, and an origin of replication, termination sequences and the like as desired.

A protein of the present invention, once expressed, can be isolated from yeast by lysing the cells and applying standard protein isolation techniques to the lysates. The monitoring of the purification process can be accomplished by using Western blot techniques or radioimmunoassay or other standard immunoassay techniques.

The proteins of the present invention can also be constructed using non-cellular synthetic methods. Solid phase synthesis of proteins of less than about 50 amino acids in length may be accomplished by attaching the C-terminal amino acid of the sequence to an insoluble support followed by sequential addition of the remaining amino acids in the sequence. Techniques for solid phase synthesis are described by Barany and Merrifield, *Solid-Phase Peptide Synthesis*, pp. 3-284 in *The Peptides*:



*Analysis, Synthesis, Biology. Vol. 2: Special Methods in Peptide Synthesis, Part A.*; Merrifield et al., *J. Am. Chem. Soc.* 85:2149-2156 (1963), and Stewart et al., *Solid Phase Peptide Synthesis, 2nd ed.*, Pierce Chem. Co., Rockford, Ill. (1984). Proteins of greater length may be synthesized by condensation of the amino and carboxy termini of shorter fragments. Methods of forming peptide bonds by activation of a carboxy terminal end (e.g., by the use of the coupling reagent N,N'-dicyclohexylcarbodiimide) is known to those of skill.

The proteins of this invention may be purified to substantial purity by standard techniques well known in the art, including detergent solubilization, selective precipitation with such substances as ammonium sulfate, column chromatography, immunopurification methods, and others. See, for instance, R. Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag: New York (1982); Deutscher, *Guide to Protein Purification*, Academic Press (1990). For example, antibodies may be raised to the proteins as described herein. Purification from *E. coli* can be achieved following procedures described in U.S. Patent No. 4,511,503. Detection of the expressed protein is achieved by methods known in the art and include, for example, radioimmunoassays, Western blotting techniques or immunoprecipitation.

The present invention further provides a method for modulating (i.e., increasing or decreasing) the concentration or composition of the polypeptides of the present invention in a plant or part thereof. Modulation of the polypeptides can be effected by increasing or decreasing the concentration and/or the composition of the polypeptides in a plant. The method comprises transforming a plant cell with an expression cassette comprising a polynucleotide of the present invention to obtain a transformed plant cell, growing the transformed plant cell under plant forming conditions, and expressing the polynucleotide in the plant for a time sufficient to modulate concentration and/or composition of the polypeptides in the plant or plant part.

In some embodiments, the content and/or composition of polypeptides of the present invention in a plant may be modulated by altering, *in vivo* or *in vitro*, the promoter of a non-isolated gene of the present invention to up- or down-regulate gene expression. In some embodiments, the coding regions of native genes of the

present invention can be altered via substitution, addition, insertion, or deletion to decrease activity of the encoded enzyme. See, e.g., Kmiec, U.S. Patent 5,565,350; Zarling et al., PCT/US93/03868.

In some embodiments, an isolated nucleic acid (e.g., a vector) comprising a promoter sequence is transfected into a plant cell. Subsequently, a plant cell comprising the isolated nucleic acid is selected for by means known to those of skill in the art such as, but not limited to, Southern blot, DNA sequencing, or PCR analysis using primers specific to the promoter and to the nucleic acid and detecting amplicons produced therefrom. A plant or plant part altered or modified by the foregoing embodiments is grown under plant forming conditions for a time sufficient to modulate the concentration and/or composition of polypeptides of the present invention in the plant. Plant forming conditions are well known in the art.

In general, concentration of the polypeptides is increased or decreased by at least 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% relative to a native control plant, plant part, or cell lacking the aforementioned expression cassette. Modulation in the present invention may occur during and/or subsequent to growth of the plant to the desired stage of development.

Modulating nucleic acid expression temporally and/or in particular tissues can be controlled by employing the appropriate promoter operably linked to a polynucleotide of the present invention in, for example, sense or antisense orientation as discussed in greater detail above. Induction of expression of a polynucleotide of the present invention can also be controlled by exogenous administration of an effective amount of inducing compound. Inducible promoters and inducing compounds that activate expression from these promoters are well known in the art.

In preferred embodiments, the polypeptides of the present invention are modulated in monocots or dicots, preferably maize, soybean, alfalfa, sunflower, *Brassica*, cotton, sorghum, wheat, barley, millet, rice, cassava, potato, *Arabidopsis*, tomato, pepper, apple, spinach, or lettuce.

Means of detecting the proteins of the present invention are not critical aspects of the present invention. In a preferred embodiment, the proteins are detected and/or quantified using any of a number of well recognized immunological binding assays

(see, e.g., U.S. Patents 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see also Methods in Cell Biology, Vol. 37: *Antibodies in Cell Biology*, Asai, Ed., Academic Press, Inc. New York (1993); *Basic and Clinical Immunology* 7th Edition, Stites & Terr, Eds. (1991). Moreover, the immunoassays of the present invention can be performed in any of several configurations, e.g., those reviewed in *Enzyme Immunoassay*, Maggio, Ed., CRC Press, Boca Raton, Florida (1980); Tijan, Practice and Theory of Enzyme Immunoassays, *Laboratory Techniques in Biochemistry and Molecular Biology*, Elsevier Science Publishers B.V., Amsterdam (1985); Harlow and Lane, *supra*; *Immunoassay: A Practical Guide*, Chan, Ed., Academic Press, Orlando, FL (1987); *Principles and Practice of Immunoassays*, Price and Newman Eds., Stockton Press, NY (1991); and *Non-isotopic Immunoassays*, Ngo, Ed., Plenum Press, NY (1988).

Typical methods for detecting proteins include Western blot (immunoblot) analysis, analytic biochemical methods such as electrophoresis, capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, and the like, and various immunological methods such as fluid or gel precipitin reactions, immunodiffusion (single or double), immunoelectrophoresis, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs), immunofluorescent assays, and the like.

Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to an anti-ligand (e.g., streptavidin) molecule that is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. A number of ligands and anti-ligands can be used. Where a ligand has a natural anti-ligand, for example, biotin, thyroxine, and cortisol, it can be used in conjunction with the labeled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody.

The molecules can also be conjugated directly to signal generating compounds, e.g., by conjugation with an enzyme or fluorophore. Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and

glycosidases, or oxidoreductases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol. For a review of various labeling or signal  
5 producing systems which may be used, see U.S. Patent No. 4,391,904, which is incorporated herein by reference.

Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples  
10 comprising the target antibodies. In this format, none of the components need be labeled and the presence of the target antibody is detected by simple visual inspection.

The proteins of the present invention can be used for identifying compounds that bind to (e.g., substrates), and/or increase or decrease (i.e., modulate) the  
15 enzymatic activity of catalytically-active polypeptides of the present invention. The method comprises contacting a polypeptide of the present invention with a compound whose ability to bind to or modulate enzyme activity is to be determined. The polypeptide employed will have at least 20%, preferably at least 30% or 40%, more preferably at least 50% or 60%, and most preferably at least 70% or 80% of the  
20 specific activity of the native, full-length polypeptide of the present invention (e.g., enzyme). Methods of measuring enzyme kinetics are well known in the art. See, e.g., Segel, *Biochemical Calculations*, 2<sup>nd</sup> ed., John Wiley and Sons, New York (1976).

Antibodies can be raised to a protein of the present invention, including  
25 individual, allelic, strain, or species variants, and fragments thereof, both in their naturally occurring (full-length) forms and in recombinant forms. Additionally, antibodies are raised to these proteins in either their native configurations or in non-native configurations. Anti-idiotypic antibodies can also be generated. Many methods of making antibodies are known to persons of skill.

30 In some instances, it is desirable to prepare monoclonal antibodies from various mammalian hosts, such as mice, rodents, primates, humans, etc. Description

of techniques for preparing such monoclonal antibodies are found in, e.g., *Basic and Clinical Immunology*, 4th ed., Stites et al., Eds., Lange Medical Publications, Los Altos, CA, and references cited therein; Harlow and Lane, *Supra*; Goding, *Monoclonal Antibodies: Principles and Practice*, 2nd ed., Academic Press, New York, NY (1986);  
5 and Kohler and Milstein, *Nature* 256:495-497 (1975).

Other suitable techniques involve selection of libraries of recombinant antibodies in phage or similar vectors (see, e.g., Huse et al., *Science* 246:1275-1281 (1989); and Ward et al., *Nature* 341:544-546 (1989); and Vaughan et al., *Nature Biotechnology*, 14:309-314 (1996)). Alternatively, high avidity human monoclonal  
10 antibodies can be obtained from transgenic mice comprising fragments of the unrearranged human heavy and light chain Ig loci (i.e., minilocus transgenic mice). Fishwild et al., *Nature Biotech.*, 14:845-851 (1996). Also, recombinant immunoglobulins may be produced. See, Cabilly, U.S. Patent No. 4,816,567; and Queen et al., *Proc. Nat'l Acad. Sci.* 86:10029-10033 (1989).

15 The antibodies of this invention can be used for affinity chromatography in isolating proteins of the present invention, for screening expression libraries for particular expression products such as normal or abnormal protein or for raising anti-idiotypic antibodies which are useful for detecting or diagnosing various pathological conditions related to the presence of the respective antigens.

20 Frequently, the proteins and antibodies of the present invention will be labeled by joining, either covalently or non-covalently, a substance which provides for a detectable signal. A wide variety of labels and conjugation techniques are known and are reported extensively in both the scientific and patent literature. Suitable labels include radionucleotides, enzymes, substrates, cofactors, inhibitors, fluorescent  
25 moieties, chemiluminescent moieties, magnetic particles, and the like.

### **Transformation of Cells**

The method of transformation/transfection is not critical to the invention; various methods of transformation or transfection are currently available. As newer  
30 methods are available to transform crops or other host cells they may be directly applied. Accordingly, a wide variety of methods have been developed to insert a

DNA sequence into the genome of a host cell to obtain the transcription and/or translation of the sequence to effect phenotypic changes in the organism. Thus, any method that provides for efficient transformation/transfection may be employed.

A DNA sequence coding for the desired polynucleotide of the present invention, for example a cDNA, RNA or a genomic sequence, will be used to construct an expression cassette that can be introduced into the desired plant. Isolated nucleic acid acids of the present invention can be introduced into plants according to techniques known in the art. Generally, expression cassettes as described above and suitable for transformation of plant cells are prepared.

Techniques for transforming a wide variety of higher plant species are well known and described in the technical, scientific, and patent literature. See, for example, Weising et al., *Ann. Rev. Genet.* 22:421-477 (1988). For example, the DNA construct may be introduced directly into the genomic DNA of the plant cell using techniques such as electroporation, PEG-mediated transfection, particle bombardment, silicon fiber delivery, or microinjection of plant cell protoplasts or embryogenic callus. See, e.g., Tomes et al., *Direct DNA Transfer into Intact Plant Cells Via Microprojectile Bombardment*. pp.197-213 in *Plant Cell, Tissue and Organ Culture, Fundamental Methods*. eds. O. L. Gamborg and G.C. Phillips. Springer-Verlag Berlin Heidelberg New York, 1995. Alternatively, the DNA constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct and adjacent marker into the plant cell DNA when the cell is infected by the bacteria. See, U.S. Patent No. 5,591,616.

The introduction of DNA constructs using polyethylene glycol precipitation is described in Paszkowski et al., *Embo J.* 3:2717-2722 (1984). Electroporation techniques are described in Fromm et al., *Proc. Natl. Acad. Sci.* 82:5824 (1985). Ballistic transformation techniques are described in Klein et al., *Nature* 327:70-73 (1987).

*Agrobacterium tumefaciens*-mediated transformation techniques are well described in the scientific literature. See, for example Horsch et al., *Science*

233:496-498 (1984), and Fraley et al., *Proc. Natl. Acad. Sci.* 80:4803 (1983). For instance, *Agrobacterium* transformation of maize is described in WO publication no. 98/32326.

Other methods of transfection or transformation include (1) *Agrobacterium*  
5 *rhizogenes*-mediated transformation (see, e.g., Lichtenstein and Fuller In: Genetic Engineering, vol. 6, PWJ Rigby, Ed., London, Academic Press, 1987; and Lichtenstein, C. P., and Draper, J., In: DNA Cloning, Vol. II, D. M. Glover, Ed., Oxford, IRI Press, 1985); Application PCT/US87/02512 (WO 88/02405 published Apr. 7, 1988) describes the use of *A. rhizogenes* strain A4 and its Ri plasmid along with *A.*  
10 *tumefaciens* vectors pARC8 or pARC16; (2) liposome-mediated DNA uptake (see, e.g., Freeman et al., *Plant Cell Physiol.* 25:1353, 1984); (3) the vortexing method (see, e.g., Kindle, *Proc. Natl. Acad. Sci. USA* 87:1228, (1990).

DNA can also be introduced into plants by direct DNA transfer into pollen as described by Zhou et al., *Methods in Enzymology*, 101:433 (1983); D. Hess, Intern  
15 Rev. Cytol., 107:367 (1987); Luo et al., *Plant Mol. Biol. Reporter*, 6:165 (1988). Expression of polypeptide coding nucleic acids can be obtained by injection of the DNA into reproductive organs of a plant as described by Pena et al., *Nature* 325:274 (1987). DNA can also be injected directly into the cells of immature embryos and the rehydration of desiccated embryos as described by Neuhaus et al., *Theor. Appl.*  
20 *Genet.*, 75:30 (1987); and Benbrook et al., in Proceedings Bio Expo 1986, Butterworth, Stoneham, Mass., pp. 27-54 (1986).

Animal and lower eukaryotic (e.g., yeast) host cells are competent or rendered competent for transfection by various means. There are several well-known methods of introducing DNA into animal cells. These include: calcium phosphate precipitation,  
25 fusion of the recipient cells with bacterial protoplasts containing the DNA, treatment of the recipient cells with liposomes containing the DNA, DEAE dextran, electroporation, biolistics, and micro-injection of the DNA directly into the cells. The transfected cells are cultured by means well known in the art. Kuchler, R.J., *Biochemical Methods in Cell Culture and Virology*, Dowden, Hutchinson and Ross, Inc. (1977).

## Transgenic Plant Regeneration

Transformed plant cells which are derived by any of the above transformation techniques can be cultured to regenerate a whole plant which possesses the transformed genotype. Such regeneration techniques often rely on manipulation of certain phytohormones in a tissue culture growth medium, typically relying on a biocide and/or herbicide marker which has been introduced together with a polynucleotide of the present invention. For transformation and regeneration of maize see, Gordon-Kamm et al., *The Plant Cell*, 2:603-618 (1990).

Plants cells transformed with a plant expression vector can be regenerated, e.g., from single cells, callus tissue or leaf discs according to standard plant tissue culture techniques. It is well known in the art that various cells, tissues, and organs from almost any plant can be successfully cultured to regenerate an entire plant. Plant regeneration from cultured protoplasts is described in Evans et al., *Protoplasts Isolation and Culture, Handbook of Plant Cell Culture*, Macmillan Publishing Company, New York, pp. 124-176 (1983); and Binding, *Regeneration of Plants, Plant Protoplasts*, CRC Press, Boca Raton, pp. 21-73 (1985).

The regeneration of plants containing the foreign gene introduced by *Agrobacterium* can be achieved as described by Horsch et al., *Science*, 227:1229-1231 (1985) and Fraley et al., *Proc. Natl. Acad. Sci. USA* 80:4803 (1983). This procedure typically produces shoots within two to four weeks and these transformant shoots are then transferred to an appropriate root-inducing medium containing the selective agent and an antibiotic to prevent bacterial growth. Transgenic plants of the present invention may be fertile or sterile.

Regeneration can also be obtained from plant callus, explants, organs, or parts thereof. Such regeneration techniques are described generally in Klee et al., *Ann. Rev. of Plant Phys.* 38:467-486 (1987). The regeneration of plants from either single plant protoplasts or various explants is well known in the art. See, for example, *Methods for Plant Molecular Biology*, A. Weissbach and H. Weissbach, eds., Academic Press, Inc., San Diego, Calif. (1988). For maize cell culture and regeneration see generally, *The Maize Handbook*, Freeling and Walbot, Eds.,



Springer, New York (1994); *Corn and Corn Improvement*, 3<sup>rd</sup> edition, Sprague and Dudley Eds., American Society of Agronomy, Madison, Wisconsin (1988).

One of skill will recognize that after the expression cassette is stably incorporated in transgenic plants and confirmed to be operable, it can be introduced  
5 into other plants by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

In vegetatively propagated crops, mature transgenic plants can be propagated by the taking of cuttings or by tissue culture techniques to produce multiple identical plants. Selection of desirable transgenics is made and new varieties are obtained  
10 and propagated vegetatively for commercial use. In seed-propagated crops, mature transgenic plants can be self crossed to produce a homozygous inbred plant. The inbred plant produces seed containing the newly introduced heterologous nucleic acid. These seeds can be grown to produce plants that would produce the selected phenotype.

15 Parts obtained from the regenerated plant, such as flowers, seeds, leaves, branches, fruit, and the like are included in the invention, provided that these parts comprise cells comprising the isolated nucleic acid of the present invention. Progeny and variants, and mutants of the regenerated plants are also included within the scope of the invention, provided that these parts comprise the introduced nucleic acid  
20 sequences.

Transgenic plants expressing a selectable marker can be screened for transmission of the nucleic acid of the present invention by, for example, standard immunoblot and DNA detection techniques. Transgenic lines are also typically evaluated on levels of expression of the heterologous nucleic acid. Expression at the  
25 RNA level can be determined initially to identify and quantitate expression-positive plants. Standard techniques for RNA analysis can be employed and include PCR amplification assays using oligonucleotide primers designed to amplify only the heterologous RNA templates and solution hybridization assays using heterologous nucleic acid-specific probes. The RNA-positive plants can then be analyzed for  
30 protein expression by Western immunoblot analysis using the specifically reactive antibodies of the present invention. In addition, *in situ* hybridization and

immunocytochemistry according to standard protocols can be done using heterologous nucleic acid specific polynucleotide probes and antibodies, respectively, to localize sites of expression within transgenic tissue. Generally, a number of transgenic lines are usually screened for the incorporated nucleic acid to identify and  
5 select plants with the most appropriate expression profiles.

A preferred embodiment is a transgenic plant that is homozygous for the added heterologous nucleic acid; i.e., a transgenic plant that contains two added nucleic acid sequences, one gene at the same locus on each chromosome of a chromosome pair. A homozygous transgenic plant can be obtained by sexually  
10 mating (selfing) a heterozygous transgenic plant that contains a single added heterologous nucleic acid, germinating some of the seed produced and analyzing the resulting plants produced for altered expression of a polynucleotide of the present invention relative to a control plant (i.e., native, non-transgenic). Back-crossing to a parental plant and out-crossing with a non- transgenic plant are also contemplated.

15 The present invention provides a method of genotyping a plant comprising a polynucleotide of the present invention. Genotyping provides a means of distinguishing homologs of a chromosome pair and can be used to differentiate segregants in a plant population. Molecular marker methods can be used for phylogenetic studies, characterizing genetic relationships among crop varieties,  
20 identifying crosses or somatic hybrids, localizing chromosomal segments affecting monogenic traits, map based cloning, and the study of quantitative inheritance. See, e.g., *Plant Molecular Biology: A Laboratory Manual*, Chapter 7, Clark, Ed., Springer-Verlag, Berlin (1997). For molecular marker methods, see generally, *The DNA Revolution* by Andrew H. Paterson 1996 (Chapter 2) in: *Genome Mapping in Plants* (ed. Andrew H. Paterson) by Academic Press/R. G. Landis Company, Austin, Texas,  
25 pp.7-21.

The particular method of genotyping in the present invention may employ any number of molecular marker analytic techniques such as, but not limited to, restriction fragment length polymorphisms (RFLPs). RFLPs are the product of allelic differences  
30 between DNA restriction fragments caused by nucleotide sequence variability. Thus, the present invention further provides a means to follow segregation of a gene or

nucleic acid of the present invention as well as chromosomal sequences genetically linked to these genes or nucleic acids using such techniques as RFLP analysis.

Plants that can be transformed in the method of the invention include monocotyledonous and dicotyledonous plants. Preferred plants include maize, soybean, alfalfa, sunflower, *Brassica*, cotton, sorghum, wheat, barley, millet, rice, cassava, potato, *Arabidopsis*, tomato, pepper, apple, spinach, and lettuce.

Seeds derived from plants regenerated from transformed plant cells, plant parts or plant tissues, or progeny derived from the regenerated transformed plants, may be used directly as feed or food, or further processing may occur.

The present invention will be further described by reference to the following detailed examples. It is understood, however, that there are many extensions, variations, and modifications on the basic theme of the present invention beyond that shown in the examples and description, which are within the spirit and scope of the present invention. All publications, patents, and patent applications cited herein are hereby incorporated by reference.

## EXAMPLES

### Construction of cDNA libraries

Total RNA is isolated from preferred plant tissues, such as roots, rhizomes, or tubers, with Promega's RNAgents (Promega, Madison, WI). In brief, tissue samples are pulverized in liquid nitrogen before the addition of a denaturing solution. Addition of a mixture of phenol and chloroform is followed by centrifugation to separate an aqueous phase and an organic phase. The total RNA is recovered by precipitation with isopropyl alcohol from the aqueous phase. Purification of mRNA is performed using the Pharmacia oligo(dT)-cellulose spun column system (Amersham Pharmacia Biotech, Uppsala, Sweden).

cDNA synthesis is performed and unidirectional cDNA libraries are constructed using products of Life Technologies Inc. (Rockville, MD) and Stratagene (La Jolla, CA). The first strand of cDNA is synthesized with SuperScript™ II Reverse Transcriptase at 37°C. Second strand synthesis is carried out at 16°C. The cDNA

fragments are purified using a sephacryl cDNA size fractionation column and ligated into an appropriate vector between EcoRI and XhoI restriction sites.

## **Sequencing of cDNA and Library Subtraction**

### 5        Sequencing Template Preparation

Individual colonies are picked and DNA prepared either by PCR with M13 forward primers and M13 reverse primers, or by plasmid isolation. All the cDNA clones are sequenced using M13 reverse primers.

### 10        Q-bot Subtraction Procedure

cDNA libraries subjected to the subtraction procedure are plated out on 22 x 22 cm<sup>2</sup> agar plate at density of about 3,000 colonies per plate. The plates are incubated in a 37°C incubator for 12-24 hours. Colonies are picked into 384-well plates by a robot colony picker, Q-bot (GENETIX Limited). These plates are  
15 incubated overnight at 37°C.

Once sufficient colonies are picked, they are pinned onto 22 x 22 cm<sup>2</sup> nylon membranes using Q-bot. Each membrane contains 9,216 colonies or 36,864 colonies. These membranes are placed onto agar plate with appropriate antibiotic. The plates are incubated at 37°C overnight.

20        After colonies are recovered on the second day, these filters are placed on filter paper prewetted with denaturing solution for four minutes, then are incubated on top of a boiling water bath for additional four minutes. The filters are then placed on filter paper prewetted with neutralizing solution for four minutes. After excess solution is removed by placing the filters on dry filter papers for one minute, the colony side of  
25 the filters are placed into Proteinase K solution, incubated at 37°C for 40-50 minutes. The filters are placed on dry filter papers to dry overnight. DNA is then cross-linked to nylon membrane by UV light treatment.

Colony hybridization is conducted as described by Sambrook, J., Fritsch, E.F. and Maniatis, T., (in Molecular Cloning: A laboratory Manual, 2nd Edition). The  
30 following probes are used in colony hybridization:

1. First strand cDNA from the same tissue as the library was made from to remove the most redundant clones.
2. 48-192 most redundant cDNA clones from the same library based on previous sequencing data.
- 5 3. 192 most redundant cDNA clones in the entire sequence database.
4. A Sal-A20 oligo nucleotide: TCG ACC CAC GCG TCC GAA AAA AAA AAA AAA AAA AAA, removes clones containing a poly A tail but no cDNA.
5. cDNA clones derived from rRNA.

10       The image of the autoradiography is scanned into computer and the signal intensity and cold colony addresses of each colony is analyzed. Re-arraying of cold-colonies from 384 well plates to 96 well plates is conducted using Q-bot.

#### **Identification of the Gene from a Computer Homology Search**

15       Gene identities were determined by conducting BLAST (Basic Local Alignment Search Tool; Altschul, S. F., et al., (1993) *J. Mol. Biol.* 215:403-410; see also [www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)) searches under default parameters for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure  
20       Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm. The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences  
25       contained in the "nr" database using the BLASTX algorithm (Gish, W. and States, D. J. *Nature Genetics* 3:266-272 (1993)) provided by the NCBI. In some cases, the sequencing data from two or more clones containing overlapping segments of DNA were used to construct contiguous DNA sequences.

## Preparation of Transgenic Plants

### Bombardment

A general method of genetic transformation used to produce transgenic maize plants is by bombardment of embryogenically responsive immature embryos with tungsten particles associated with DNA plasmids, said plasmids consisting of a selectable and an unselectable marker gene.

### Preparation of Tissue

Immature embryos of "High Type II" are the target for particle bombardment-mediated transformation. This genotype is the F<sub>1</sub> of two purebred genetic lines, parent A and parent B, derived from A188 X B73. Both parents are selected for high competence of somatic embryogenesis. See Armstrong et al., "Development and Availability of Germplasm with High Type II Culture Formation Response," *Maize Genetics Cooperation Newsletter*, Vol. 65, pp. 92 (1991); incorporated herein in its entirety by reference.

Ears from F<sub>1</sub> plants are selfed or sibbed, and embryos are aseptically dissected from developing caryopses when the scutellum first becomes opaque. The proper stage occurs about 9-13 days post-pollination, and most generally about 10 days post-pollination, and depends on growth conditions. The embryos are about 0.75 to 1.5 mm long. Ears are surface sterilized with 20-50% Clorox for 30 min, followed by 3 rinses with sterile distilled water.

Immature embryos are cultured, scutellum oriented upward, on embryogenic induction medium comprised of N6 basal salts (Chu et al., "Establishment of an Efficient Medium for Anther Culture of Rice Through Comparative Experiments on the Nitrogen Sources," *Scientia Sinica*, (Peking), Vol. 18, pp. 659-668 (1975); incorporated herein in its entirety by reference; Eriksson vitamins (See Eriksson, T., "Studies on the Growth Requirements and Growth Measurements of *Haplopappus gracilis*," *Physiol. Plant*, Vol. 18, pp. 976-993 (1965); incorporated herein in its entirety by reference), 0.5 mg/l thiamine HCl, 30 gm/l sucrose, 2.88 gm/l L-proline, 1 mg/l 2,4-dichlorophenoxyacetic acid, 2 gm/l Gelrite, and 8.5 mg/l AgNO<sub>3</sub>.

The medium is sterilized by autoclaving at 121°C for 15 min and dispensed into 100 X 25 mm petri dishes. AgNO<sub>3</sub> is filter-sterilized and added to the medium

after autoclaving. The tissues are cultured in complete darkness at 28°C. After about 3 to 7 days, generally about 4 days, the scutellum of the embryo has swelled to about double its original size and the protuberances at the coleorhizal surface of the scutellum indicate the inception of embryogenic tissue. Up to 100% of the embryos  
5 display this response, but most commonly, the embryogenic response frequency is about 80%.

When the embryogenic response is observed, the embryos are transferred to a medium comprised of induction medium modified to contain 120 gm/l sucrose. The embryos are oriented with the coleorhizal pole, the embryogenically responsive  
10 tissue, upwards from the culture medium. Ten embryos per petri dish are located in the center of a petri dish in an area about 2 cm in diameter. The embryos are maintained on this medium for 3-16 hr, preferably 4 hours, in complete darkness at 28°C just prior to bombardment with particles associated with plasmid DNAs containing the selectable and unselectable marker genes.

15 To effect particle bombardment of embryos, the particle-DNA agglomerates are accelerated using a DuPont PDS-1000 particle acceleration device. The particle-DNA agglomeration is briefly sonicated and 10 µl are deposited on macrocarriers and the ethanol allowed to evaporate. The macrocarrier is accelerated onto a stainless-steel stopping screen by the rupture of a polymer diaphragm (rupture disk). Rupture  
20 is effected by pressurized helium. Depending on the rupture disk breaking pressure, the velocity of particle-DNA acceleration may be varied. Rupture disk pressures of 200 to 1800 psi are commonly used, with those of 650 to 1100 psi being more preferred, and about 900 psi being most highly preferred. Rupture disk breaking pressures are additive so multiple disks may be used to effect a range of rupture  
25 pressures.

Preferably, the shelf containing the plate with embryos is 5.1 cm below the bottom of the macrocarrier platform (shelf #3), but may be located at other distances. To effect particle bombardment of cultured immature embryos, a rupture disk and a macrocarrier with dried particle-DNA agglomerates are installed in the device. The  
30 He pressure delivered to the device is adjusted to 200 psi above the rupture disk breaking pressure. A petri dish with the target embryos is placed into the vacuum

chamber and located in the projected path of accelerated particles. A vacuum is created in the chamber, preferably about 28 inches Hg. After operation of the device, the vacuum is released and the petri dish is removed.

Bombarded embryos remain on the osmotically adjusted medium during bombardment, and preferably for two days subsequently, although the embryos may remain on this medium for 1 to 4 days. The embryos are transferred to selection medium comprised of N6 basal salts, Eriksson vitamins, 0.5 mg/l thiamine HCl, 30 gm/l sucrose, 1 mg/l 2,4-dichlorophenoxyacetic acid, 2 gm/l Gelrite, 0.85 mg/l AgNO<sub>3</sub> and 3 mg/l bialaphos. Bialaphos is added filter-sterilized. The embryos are subcultured to fresh selection medium at 10 to 14 day intervals. After about 7 weeks, embryogenic tissue, putatively transgenic for both selectable and unselectable marker genes, is seen to proliferate from about 7% of the bombarded embryos. Putative transgenic tissue is rescued, and that tissue derived from individual embryos is considered to be an event and is propagated independently on selection medium. Two cycles of clonal propagation are achieved by visual selection for the smallest contiguous fragments of organized embryogenic tissue.

For regeneration of transgenic plants, embryogenic tissue is subcultured to medium comprised of MS salts and vitamins (Murashige, T. and F. Skoog, "A revised medium for rapid growth and bio assays with tobacco tissue cultures," *Physiologia Plantarum* 15:473-497 (1962), incorporated herein in its entirety by reference), 100 mg/l myo-inositol, 60 gm/l sucrose, 3 gm/l Gelrite, 0.5 mg/l zeatin, 1 mg/l indole-3-acetic acid, 26.4 ng/l cis-trans-abscissic acid, and 3 mg/l bialaphos in 100 X 25 mm petri dishes and incubated in darkness at 28°C until the development of well-formed, matured somatic embryos can be visualized. This requires about 14 days.

Well-formed somatic embryos are opaque and cream-colored, and are comprised of an identifiable scutellum and coleoptile. The embryos are individually subcultured to germination medium comprised of MS salts and vitamins, 100 mg/l myo-inositol, 40 gm/l sucrose and 1.5 gm/l Gelrite in 100 X 25 mm petri dishes and incubated under a 16 hr light: 8 hr dark photoperiod and 40  $\mu\text{Einsteinsm}^{-2}\text{sec}^{-1}$  from cool-white fluorescent tubes. After about 7 days, the somatic embryos have germinated and produced a well-defined shoot and root. The individual plants are



subcultured to germination medium in 125 x 25 mm glass tubes to allow further plant development. The plants are maintained under a 16 hr light: 8 hr dark photoperiod and  $40 \mu\text{Einsteinsm}^{-2}\text{sec}^{-1}$  from cool-white fluorescent tubes.

After about 7 days, the plants are well-established and are transplanted to horticultural soil, hardened off, and potted into commercial greenhouse soil mixture and grown to sexual maturity in a greenhouse. An elite inbred line is used as a male to pollinate regenerated transgenic plants.

#### Preparation of Particles

Fifteen mg of tungsten particles (General Electric), 0.5 to 1.8  $\mu\text{m}$ , preferably 1 to 1.8  $\mu\text{m}$ , and most preferably 1  $\mu\text{m}$ , are added to 2 ml of concentrated nitric acid. This suspension is sonicated at  $0^\circ\text{C}$  for 20 min (Branson Sonifier Model 450, 40% output, constant duty cycle). Tungsten particles are pelleted by centrifugation at 10,000 rpm (Biofuge) for 1 min and the supernatant is removed. Two ml of sterile distilled water are added to the pellet and sonicated briefly to resuspend the particles. The suspension is pelleted, 1 ml of absolute ethanol is added to the pellet and sonicated briefly to resuspend the particles. The particles are rinsed, pelleted, and resuspended a further 2 times with sterile distilled water, and finally the particles are resuspended in 2 ml of sterile distilled water. The particles are subdivided into 250  $\mu\text{l}$  aliquots and stored frozen.

#### Preparation of particle-plasmid DNA association

The stock of tungsten particles is sonicated briefly in a water bath sonicator (Branson Sonifier Model 450, 20% output, constant duty cycle) and 50  $\mu\text{l}$  is transferred to a microfuge tube. Plasmid DNA is added to the particles for a final DNA amount of 0.1 to 10  $\mu\text{g}$  in 10  $\mu\text{l}$  total volume, and briefly sonicated. Preferably 1  $\mu\text{g}$  total DNA is used. Specifically, 5  $\mu\text{l}$  of PHP8001 (gz::HT12::gz) and 5  $\mu\text{l}$  of PHP3528 (enhanced CAMV::Bar::PinII), at 0.1  $\mu\text{g}/\mu\text{l}$  in TE buffer, are added to the particle suspension. Fifty  $\mu\text{l}$  of sterile aqueous 2.5 M  $\text{CaCl}_2$  are added, and the mixture is briefly sonicated and vortexed. Twenty  $\mu\text{l}$  of sterile aqueous 0.1M spermidine are added and the mixture is briefly sonicated and vortexed. The mixture is incubated at room temperature for 20 min with intermittent brief sonication. The

particle suspension is centrifuged, and the supernatant is removed. Two hundred fifty  $\mu$ l of absolute ethanol is added to the pellet and briefly sonicated. The suspension is pelleted, the supernatant is removed, and 60  $\mu$ l of absolute ethanol is added. The suspension is sonicated briefly before loading the particle-DNA agglomeration onto  
5 macrocarriers.

### ***Agrobacterium* co-cultivation**

Another method of transformation is by co-cultivation with *Agrobacterium*.

*Agrobacterium* is streaked out from a -80° frozen aliquot onto a plate containing PHI-L medium and cultured at 28°C in the dark for 3 days. PHI-L media comprises 25 ml/l  
10 Stock Solution A, 25 ml/l Stock Solution B, 450.9 ml/l Stock Solution C and spectinomycin (Sigma Chemicals) added to a concentration of 50 mg/l in sterile ddH<sub>2</sub>O (stock solution A: K<sub>2</sub>HPO<sub>4</sub> 60.0 g/l, NaH<sub>2</sub>PO<sub>4</sub> 20.0 g/l, adjust pH to 7.0 w/KOH and autoclave; stock solution B: NH<sub>4</sub>Cl 20.0 g/l, MgSO<sub>4</sub>.7H<sub>2</sub>O 6.0 g/l, KCl 3.0  
15 g/l, CaCl<sub>2</sub> 0.20 g/l, FeSO<sub>4</sub>.7H<sub>2</sub>O 50.0 mg/l, autoclave; stock solution C: glucose 5.56g/l, agar 16.67 g/l (#A-7049, Sigma Chemicals, St. Louis, MO) and autoclave).

The plate can be stored at 4°C and used usually for about 1 month. A single colony is picked from the master plate and streaked onto a plate containing PHI-M medium [yeast extract (Difco) 5.0 g/l; peptone (Difco) 10.0 g/l; NaCl 5.0 g/l; agar  
20 (Difco) 15.0 g/l; pH 6.8, containing 50 mg/L spectinomycin] and incubated at 28°C in the dark for 2 days. Five ml of either PHI-A, [CHU(N6) basal salts (Sigma C-1416) 4.0 g/l, Eriksson's vitamin mix (1000X, Sigma-1511) 1.0 ml/l; thiamine.HCl 0.5 mg/l (Sigma); 2,4-dichlorophenoxyacetic acid (2,4-D, Sigma) 1.5 mg/l; L-proline (Sigma) 0.69 g/l; sucrose (Mallinckrodt) 68.5 g/l; glucose (Mallinckrodt) 36.0 g/l; pH 5.2] for the  
25 PHI basic medium system, or PHI-I [MS salts (GIBCO BRL) 4.3 g/l; nicotinic acid (Sigma) 0.5 mg/l; pyridoxine.HCl (Sigma) 0.5 mg/l; thiamine.HCl 1.0 mg/l; myo-inositol (Sigma) 0.10 g/l; vitamin assay casamino acids (Difco Lab) 1.0 g/l; 2, 4-D 1.5 mg/l; sucrose 68.50 g/l; glucose 36.0 g/l; adjust pH to 5.2 w/KOH and filter-sterilize] for the PHI combined medium system and 5  $\mu$ l of 100 mM (3'-5'-Dimethoxy-4'-  
30 hydroxyacetophenone, Aldrich chemicals) are added to a 14 ml Falcon tube in a hood. About 3 full loops (5 mm loop size) *Agrobacterium* is collected from the plate

and suspended in the tube, then the tube is vortexed to make an even suspension. One ml of the suspension is transferred to a spectrophotometer tube and the OD of the suspension adjusted to 0.72 at 550 nm by adding either more *Agrobacterium* or more of the same suspension medium, for an *Agrobacterium* concentration of approximately  $0.5 \times 10^9$  cfu/ml to  $1 \times 10^9$  cfu/ml. The final *Agrobacterium* suspension is aliquoted into 2 ml microcentrifuge tubes, each containing 1 ml of the suspension. The suspensions are then used as soon as possible.

Embryo isolation, infection and co-cultivation:

About 2 ml of the same medium (here PHI-A or PHI-I) used for the *Agrobacterium* suspension are added into a 2 ml microcentrifuge tube. Immature embryos are isolated from a sterilized ear with a sterile spatula (Baxter Scientific Products S1565) and dropped directly into the medium in the tube. A total of about 100 embryos are placed in the tube. The optimal size of the embryos is about 1.0-1.2 mm. The cap is then closed on the tube and the tube vortexed with a Vortex Mixer (Baxter Scientific Products S8223-1) for 5 sec. at maximum speed. The medium is removed and 2 ml of fresh medium are added and the vortexing repeated. All of the medium is drawn off and 1 ml of *Agrobacterium* suspension is added to the embryos and the tube vortexed for 30 sec. The tube is allowed to stand for 5 min. in the hood. The suspension of *Agrobacterium* and embryos was poured into a Petri plate containing either PHI-B medium [CHU(N6) basal salts (Sigma C-1416) 4.0 g/l; Eriksson's vitamin mix (1000X, Sigma-1511) 1.0 ml/l; thiamine.HCl 0.5 mg/l; 2,4-D 1.5 mg/l; L-proline 0.69 g/l; silver nitrate 0.85 mg/l; gelrite (Sigma) 3.0 g/l; sucrose 30.0 g/l; acetosyringone 100  $\mu$ M; pH 5.8], for the PHI basic medium system, or PHI-J medium [MS Salts 4.3 g/l; nicotinic acid 0.50 mg/l; pyridoxine HCl 0.50 mg/l; thiamine.HCl 1.0 mg/l; myo-inositol 100.0 mg/l; 2, 4-D 1.5 mg/l; sucrose 20.0 g/l; glucose 10.0 g/l; L-proline 0.70 g/l; MES (Sigma) 0.50 g/l; 8.0 g/l agar (Sigma A-7049, purified) and 100  $\mu$ M acetosyringone with a final pH of 5.8 for the PHI combined medium system. Any embryos left in the tube are transferred to the plate using a sterile spatula. The *Agrobacterium* suspension is drawn off and the embryos placed axis side down on the media. The plate is sealed with Parafilm tape or Pylon Vegetative Combine Tape (product named "E.G.CUT" and is available in 18 mm x 50

m sections; Kyowa Ltd., Japan) and incubated in the dark at 23-25°C for about 3 days of co-cultivation.

Resting, selection and regeneration steps:

For the resting step, all of the embryos are transferred to a new plate  
5 containing PHI-C medium [CHU(N6) basal salts (Sigma C-1416) 4.0 g/l; Eriksson's vitamin mix (1000X Sigma-1511) 1.0 ml/l; thiamine.HCl 0.5 mg/l; 2.4-D 1.5 mg/l; L-proline 0.69 g/l; sucrose 30.0 g/l; MES buffer (Sigma) 0.5 g/l; agar (Sigma A-7049, purified) 8.0 g/l; silver nitrate 0.85 mg/l; carbenicillin 100 mg/l; pH 5.8]. The plate is sealed with Parafilm or Pylon tape and incubated in the dark at 28°C for 3-5 days.

10 Longer co-cultivation periods may compensate for the absence of a resting step since the resting step, like the co-cultivation step, provides a period of time for the embryo to be cultured in the absence of a selective agent. Those of ordinary skill in the art can readily test combinations of co-cultivation and resting times to optimize or improve the transformation frequency of other inbreds without undue  
15 experimentation.

For selection, all of the embryos are then transferred from the PHI-C medium to new plates containing PHI-D medium, as a selection medium, [CHU(N6) basal salts (SIGMA C-1416) 4.0 g/l; Eriksson's vitamin mix (1000X, Sigma-1511) 1.0 ml/l; thiamine.HCl 0.5 mg/l; 2.4-D 1.5 mg/l; L-proline 0.69 g/l; sucrose 30.0 g/l; MES buffer  
20 0.5 g/l; agar (Sigma A-7049, purified) 8.0 g/l; silver nitrate 0.85 mg/l; carbenicillin (ICN, Costa Mesa, CA) 100 mg/l; bialaphos (Meiji Seika K.K., Tokyo, Japan) 1.5 mg/l for the first two weeks followed by 3 mg/l for the remainder of the time.; pH 5.8] putting about 20 embryos onto each plate. The plates are sealed as described above and incubated in the dark at 28°C for the first two weeks of selection. The embryos  
25 are transferred to fresh selection medium at two-week intervals. The tissue is subcultured by transferring to fresh selection medium for a total of about 2 months. The herbicide-resistant calli are then "bulked up" by growing on the same medium for another two weeks until the diameter of the calli is about 1.5-2 cm.

For regeneration, the calli are then cultured on PHI-E medium [MS salts 4.3  
30 g/l; myo-inositol 0.1 g/l; nicotinic acid 0.5 mg/l, thiamine.HCl 0.1 mg/l, Pyridoxine.HCl 0.5 mg/l, Glycine 2.0 mg/l, Zeatin 0.5 mg/l, sucrose 60.0 g/l, Agar (Sigma, A-7049)

8.0 g/l, Indoleacetic acid (IAA, Sigma) 1.0 mg/l, Absciscic acid (ABA, Sigma) 0.1  $\mu$ M, Bialaphos 3 mg/l, carbenicillin 100 mg/l adjusted to pH 5.6] in the dark at 28°C for 1-3 weeks to allow somatic embryos to mature. The calli are then cultured on PHI-F medium (MS salts 4.3 g/l; myo-inositol 0.1 g/l; Thiamine.HCl 0.1 mg/l, Pyridoxine.HCl 0.5 mg/l, Glycine 2.0 mg/l, nicotinic acid 0.5 mg/l; sucrose 40.0 g/l; gelrite 1.5 g/l; pH 5.6] at 25°C under a daylight schedule of 16 hrs. light (270  $\mu$ E m<sup>-2</sup>sec<sup>-1</sup>) and 8 hrs. dark until shoots and roots develop. Each small plantlet is then transferred to a 25x150 mm tube containing PHI-F medium and grown under the same conditions for approximately another week. The plants are transplanted to pots with soil mixture in a greenhouse. GUS+ events are determined at the callus stage or regenerated plant stage.

For Hi-II a preferred optimized protocol was  $0.5 \times 10^9$  cfu/ml *Agrobacterium* (Table 6), a 3-5 day resting step (Example 5), and no AgNO<sub>3</sub> in the infection medium (PHI-A medium). The examples provide a variety of experiments that similarly teach those of ordinary skill in the art to optimize transformation frequencies for other maize lines.

It will be appreciated by those skilled in the art that while the invention has been described above in connection with particular embodiments and examples, the invention is not necessarily so limited and that numerous other embodiments, examples, uses, modifications and departures from the embodiments, examples and uses may be made without departing from the inventive scope of this application.